

CUAHSI Services to Enable Data Fusion

Tony Castronova acastronova@cuahsi.org

Hydrologic Scientist

CUAHSI

04-23-2018





CUAHSI in a Nutshell

- CUAHSI is a 501(c)3 Non-Profit Consortium of about 130 U.S. Academic Institutions, Non-Profits, and International Universities; also open to private organizations
- CUAHSI's mission is to shape the future of hydrologic science by:
 - Strengthening collaboration
 - Developing and delivering data, models, instrumentation and technologies
 - Promoting education
- Key Activities
 - **Community Services**, such as workshops, community meetings, training, etc.
 - **Data and Model Services** software to support science, including HIS and HydroShare





Community Services



Early Career Services

Resources and training to build capacity and extend capabilities

- Learn a new instrumentation method
- Engage your local community on water issues
- Develop or improve data tools
- Add a new field site to existing research
- The Summer Innovators program - contribute to the National Water Center goals
- Collaborate with researchers at other institutions



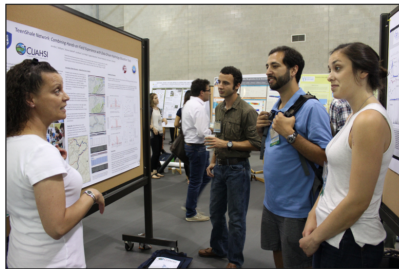


Community Meetings

2018 CUAHSI Biennial Colloquium

July 29 – August 2 at the National Conservation Training Center in Shepherdstown, WV

- Hydrologic Connections: Climate, Food, Energy, Environment, and Society
- Workshops on drones and NASA SWOT mission
- Sessions on water and climate, hydrologic feedbacks, energy-water systems, etc.





Summer Innovators Program

Training the next-generation of water scientists



- Partnership between CUAHSI, NWC, and the NWS
- Involvement of academic community, NCAR, DHS, US ACE, USGS
- Engages the academic research community in the enhancement of the National Water Model
- Fourth summer institute emphasis on groundwater and channel processes



Summer Innovators Program Outcomes and broader engagement

- More than 100 students trained on NWM creating next generation of researchers and users, more than half of their advisors engaged in the process
- Advances in the NWM functionality and improved understanding of performance
- Growing engagement of academic community
- JAWRA Featured collections and CUAHSI – NOAA Technical Reports
- Transformational educational experience

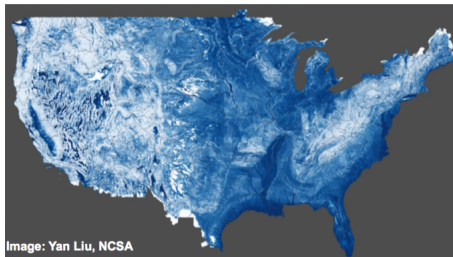
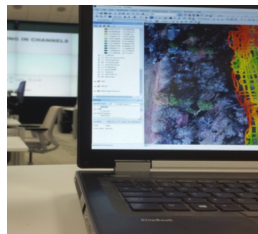


Image: Yan Liu, NCSA



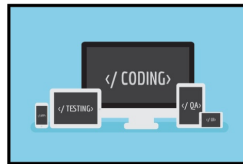
Data Services



Data Services

Resources to access data and models for research and education

- Tools to enable science
- Share, discover, analyze, collaborate
- Free and open-source software
- Fulfill data management plans
- Research and classroom activities
- Driven by community needs





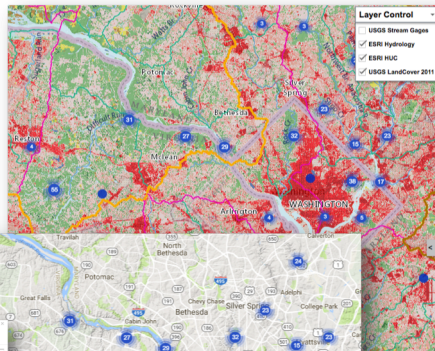
-
- The diagram illustrates the ODM CV Web Services architecture. A central box labeled "ODM CV Web Services" is connected by blue arrows to ten blue cylinders labeled "ODM" distributed across a map of the United States. A green cylinder labeled "ODM CV Central Database" is also connected to the central box.



HydroClient

A CUAHSI web client for the HIS - <http://data.cuahsi.org>

- Discover, view, download
- Federal, university, community data
- Enhancements to search and discovery interface and backend cataloging via Solr
- Integration with HydroShare Apps



Search Results in: Selected Area
Table filters are applied to map

Search:

Publisher	Service Title	Keyword	Site Name	Data Type	Value Type	Sample Medium	QC Level	Method	Collector
U.S. Geological Sur...	▼ NWIS Daily V...	Discharge, stream	PANX BRANCH NE...	average	Derived Value	SurfaceWater	Mixed	No method specified	U.S. Geological Sur...
U.S. Geological Sur...	▼ NWIS Daily V...	Discharge, stream	PEACOCK CREEK...	average	Derived Value	SurfaceWater	Mixed	No method specified	U.S. Geological Sur...
U.S. Geological Sur...	▼ NWIS Daily V...	Discharge, stream	NORTHEAST BR A...	average	Derived Value	SurfaceWater	Mixed	No method specified	U.S. Geological Sur...
U.S. Geological Sur...	▼ NWIS Daily V...	Discharge, stream	NORTHEAST BAWN...	average	Derived Value	SurfaceWater	Mixed	No method specified	U.S. Geological Sur...
U.S. Geological Sur...	▼ NWIS Daily V...	Discharge, stream	TUNNEY BRANCH...	average	Derived Value	SurfaceWater	Mixed	No method specified	U.S. Geological Sur...
U.S. Geological Sur...	▼ NWIS Daily V...	Discharge, stream	BLOOD CREEK MA...	average	Derived Value	SurfaceWater	Mixed	No method specified	U.S. Geological Sur...
U.S. Geological Sur...	▼ NWIS Daily V...	Discharge, stream	SWITS BRANCH AT...	average	Derived Value	SurfaceWater	Mixed	No method specified	U.S. Geological Sur...
U.S. Geological Sur...	▼ NWIS Daily V...	Discharge, stream	ROCK CREEK AT JO...	average	Derived Value	SurfaceWater	Mixed	No method specified	U.S. Geological Sur...
U.S. Geological Sur...	▼ NWIS Daily V...	Discharge, stream	HONEY RUN AT LA...	average	Derived Value	SurfaceWater	Mixed	No method specified	U.S. Geological Sur...

Showing 1 to 10 of 487 entries

By Publisher: By Service Title: By Keyword: By Site Name: By Data Type: By Value Type: By Sample Medium: QC Level: Method: Collector:

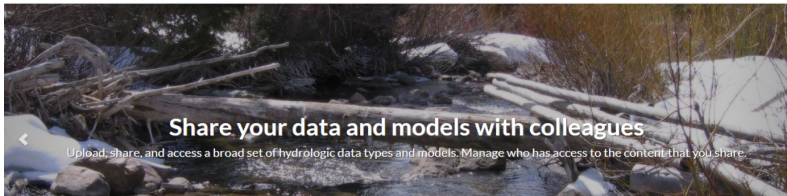
Show: entries

Previous: 1 2 3 4 5 Next:



HydroShare

Dealing with data disparity

[MY RESOURCES](#)[DISCOVER](#)[APPS](#)[HELP](#)[SIGN IN](#)

Advancing Hydrologic Understanding

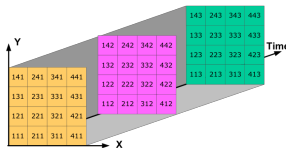
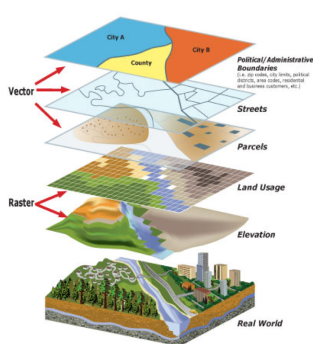
- requires integration of information from multiple sources
- may be data and computationally intensive
- requires collaboration and working as a team/community



HydroShare

Dealing with data disparity

- Time series
- Geographic raster
- Geographic feature
- Multidimensional space/time
- Model programs
- Model instances
- Combinations of the above



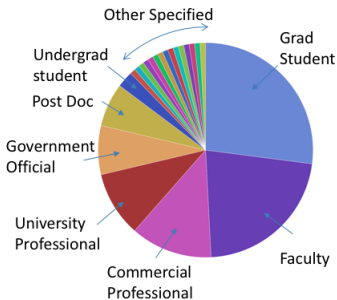
adapted from Jeff Horsburgh



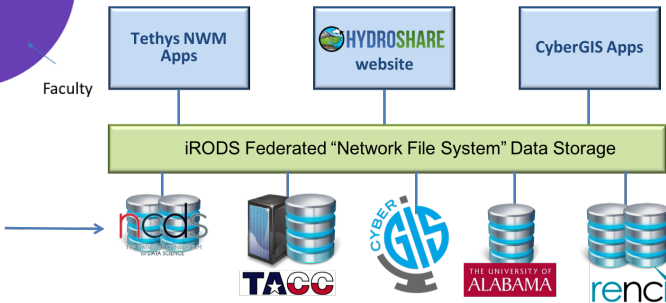
HydroShare

An architectural overview

- 50 TB rolling window of NWM forecasts (2-3 TB/day)
- National HAND layer
- 50 TB store for other HydroShare data



NOMADS
The NOAA National Operational
Model Archive and Distribution System





HydroShare

An overview of key functionality

HYDROSHARE MY RESOURCES DISCOVER COLLABORATE APPS HELP

My Resources

+ Create new

Filter

- ☐ Owned by me 50
- ☐ Editable by me 38
- ☐ Viewable by me 63
- ☐ Favorites 0

Labels

Type	Title
	Great Salt Lake Level at
	Collection of workshop the CUAHSI biennial symposium
	Material for HydroShare Biennial Symposium

Sharing status:

- Public
- Shareable

Content

contents

- GSLM_Bathymetry_noProposedPonds.csv
- GSLM_Volume_Calcs_20110915_noProposedPonds.csv
- BathymetryNotes.docx

Authors

The people or organizations that created the intellectual content of the resource.

Name	Organization	Address
David Tarboton	Utah State University	4110 Old Main Hill, Logan, UT 84322

Contributors

People or organizations that contributed technically, materially, financially, or provided general support for the creation of the resource.

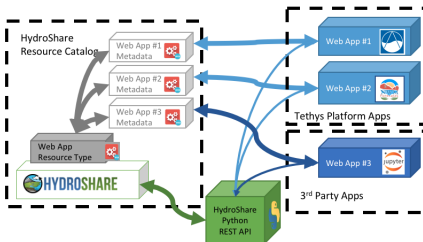
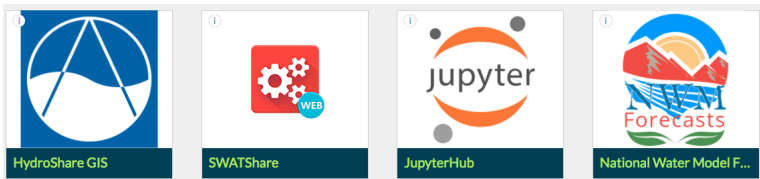
Name	Organization	Address
Blaise Chanson	BIO-WEST	1063 West 1400 North Logan, Utah 84321-2299
Glen Busch	BIO-WEST	1063 West 1400 North Logan, Utah 84321-2299

- Upload data to HydroShare
- Enhance scientific value with metadata
- Manage access and control
- Collaborate
- Publish



HydroShare

Interacting with data



Contents of /nwm/harvey/nwm.20170818/medium_range

Name	Last Modified	Size	
nwm.t00x.medium_range.channel_rt.f003.conus.nc	2017-08-18T06:13:01	3644894	dl
nwm.t00x.medium_range.channel_rt.f006.conus.nc	2017-08-18T06:13:02	3623176	dl
nwm.t00x.medium_range.channel_rt.f009.conus.nc	2017-08-18T06:12:52	3627995	dl
nwm.t00x.medium_range.channel_rt.f012.conus.nc	2017-08-18T06:12:54	3616857	dl
nwm.t00x.medium_range.channel_rt.f015.conus.nc	2017-08-18T06:12:54	3608756	dl
nwm.t00x.medium_range.channel_rt.f018.conus.nc	2017-08-18T06:12:57	3595279	dl
nwm.t00x.medium_range.channel_rt.f021.conus.nc	2017-08-18T06:12:54	3407687	dl
nwm.t00x.medium_range.channel_rt.f024.conus.nc	2017-08-18T06:12:54	3630836	dl
nwm.t00x.medium_range.channel_rt.f027.conus.nc	2017-08-18T06:13:01	3665497	dl
nwm.t00x.medium_range.channel_rt.f030.conus.nc	2017-08-18T06:12:53	3654165	dl
nwm.t00x.medium_range.channel_rt.f033.conus.nc	2017-08-18T06:12:46	3431127	dl
nwm.t00x.medium_range.channel_rt.f036.conus.nc	2017-08-18T06:12:57	3418929	dl



National Water Model Viewer

Access to daily forecasts

Home

Subset

Add Watershed

Model Configuration

Medium Range

Model Output File

Channel

Variable

Streamflow

COMID

18229909

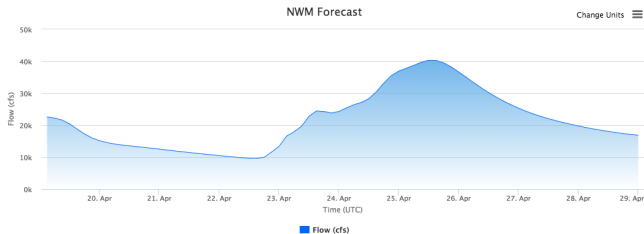
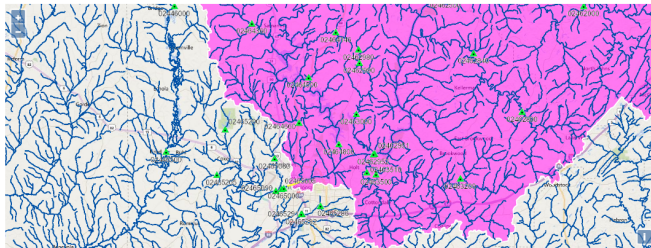
Begin Date

2018-04-19

Model Initialization Time (UTC)

00:00

View Forecast





A cloud-based coding environment - <https://jupyter.cuahsi.org/>

The screenshot displays the HydroShare Python Notebook Server interface. At the top, there's a menu bar with options like File, Edit, View, Insert, Cell, Kernel, Widgets, and Help. Below it is a toolbar with icons for various notebook actions.

Welcome to the HydroShare Python Notebook Server

You have arrived at the HydroShare Python Notebook Server landing page. The cells below demonstrate HydroShare content inside this Python scripting environment. You will be asked to provide your HydroS the notebook server can establish a secure connection with your resources. When accessing resources notebook files (i.e. *.ipynb), you will be provided hyperlinks to launch them in separate windows. To make changes from this session back into HydroShare using the `hs_utils` library.

Jupyter notebook servers enable scientists to explore, modify, and interact with data using the Python programming language. A **Jupyter notebook** is an enhanced code editor that combines text and code execution into a single script-like container. The HydroShare Python Notebook Server provides standard notebook functionality with the HydroShare data hosting platform to provide a seamless experience for water scientists.

See what scientists are doing with notebooks:

- [Raster Processing using TauDEM](#)
- [LandLab Landslide Calculations](#)
- [Basic Time Series Analysis - GenericResource](#)
- [Basic Time Series Analysis - TimeSeriesResource](#)
- [RHESSys Modeling](#)

```
# define a function to download nwm results
def get_nwm(q, iolock, out_g, outdir, cnt):
    while True:
        resid = q.get()
        if resid is None:
            break
        liget ~f {resid} {outdir}

        # get the NWM channel_rt forecasts in parallel using iRODS
        NCORE = 40
        in_q = mp.Queue(maxsize=NCORE)
        out_g = mp.Queue()
        cnt = mp.Value('i', 0)
        iolock = mp.Lock()

        pool = mp.Pool(NCORE, initializer=get_nwm,
                        initargs=(in_q, iolock, out_g, nwm_dir, cnt))

        for f in md_range:
```

feature_id	reference_time	time	streamflow	nudge	q_lateral	velocity
101	2017-07-18	2017-07-18 03:00:00	0.90	0.0	0.0	0.03
		2017-07-18 06:00:00	0.95	0.0	0.0	0.03
		2017-07-18 09:00:00	1.00	0.0	0.0	0.03

	streamflow	nudge	q_lateral	velocity
count	2.154210e+08	2.173518e+08	2.173518e+08	2.154210e+08
mean	8.026310e+00	1.104753e-05	5.058141e-03	3.091742e-02
std	2.037371e+02	6.757011e-02	1.913766e-01	1.461884e-01

NWM Results by COMID



JupyterHub

Supporting scientific reproducibility

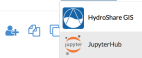
- Reproducible Science
- Scientific Collaboration
- Education
- Computationally Intensive Research
- Data Intensive Research



MY RESOURCES DISCOVER COLLABORATE APPS HELP ABOUT

SUMMA - Celia 1990 Simulation Notebook

Open with...



Authors: Anthony Castronova
Owners: Anthony Castronova
Resource type: Composite Resource
Created: Mar 16, 2018 at 7:06 p.m.
Last updated: Mar 16, 2018 at 7:07 p.m. by Anthony Castronova

 summa-celia-1990-simulation.ipynb	19.6 KB	GenericLogi...ipynb File
 pythonlibs.zip	4.0 MB	GenericLogi...zip File

Import Libraries and Prepare Environment

```
In [1]: !import os, sys

# unzip the python libraries
!unzip -oq pythonlibs.zip

# add our libs directory to the path
sys.path.append('../pythonlibs')
sys.path.append('../pythonlibs/lib/python3.6/site-packages')
sys.path.append('../pythonlibs/bin')

In [2]: #import celeryworker as c
from pyresim.Simulation import Simulation
from hydroshare import hydroshare
```

Collect Input Data from HydroShare

- Connect to HydroShare
- Download simulation input data

```
In [3]: hs = hydroshare.hydroshare()

Please enter your HydroShare username: 'TonyCastronova'
Enter the HydroShare password for user 'TonyCastronova': .....
Successfully established a connection with HydroShare

In [4]: hs.get_resource_from_hydroshare('6686dc8f26ac4bce8ff4c11e5953a74c')

Download Finished
Successfully downloaded resource 6686dc8f26ac4bce8ff4c11e5953a74c

Found the following file(s) associated with this HydroShare resource.

celia-simulation.zip

These files are stored in a dictionary called hs.content for your convenience. To access a file,
simply issue the following command where MY_FILE is one of the files listed above:

hs.content[MY_FILE]

Look at the content of the HydroShare resource that was downloaded

In [5]: print(hs.content)
data_archive = hs.content['celia-simulation.zip']
```



Meeting the Needs of the Community

- CUAHSI is model agnostic
- Provide community tools such as HIS and HydroShare, along with documentation and API's
- Provide community data sets for selected use cases through CUAHSI Water Data Services. (Harvey/Irma and Maria RAPID awards)
- Continue with Summer Institute and potentially similar activities
- Organize and host community meetings



Thank You

Tony Castronova
acastronova@cuahsi